

AD-A135 242 ON THE EQUIVALENCE OF PROBABILITY MEASURES(U) NORTH  
CAROLINA UNIV AT CHAPEL HILL DEPT OF STATISTICS  
C R BAKER ET AL. 1981 N00014-75-C-0491

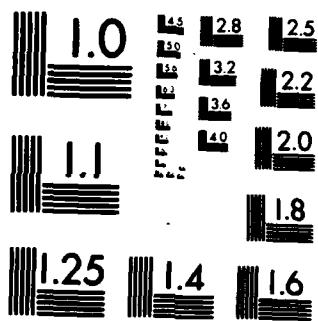
1/1

UNCLASSIFIED

F/G 12/1

NL

END  
DATE FILMED  
12.83  
DTIC



MICROCOPY RESOLUTION TEST CHART  
NATIONAL BUREAU OF STANDARDS-1963-A

A D - A 2 3 5 2 4 2

(4)

## ON THE EQUIVALENCE OF PROBABILITY MEASURES

C. R. Baker,\* A. F. Gualtierotti\*\*

\*Department of Statistics, University of North Carolina, Chapel Hill,  
NC, 27514.

\*\*IDHEAP, BFSH 1, University of Lausanne; Lausanne, Switzerland 1015

This document has been approved  
for public release and sale; its  
distribution is unlimited.

Research supported under ONR Contracts N00014-75-C-0491 and  
N00014-81-K-0373.

DTIC FILE COPY

DTIC  
SELECTED  
DEC 01 1983  
S E D  
E

This document has been approved  
for public release and sale; its  
distribution is unlimited.

83 12 01 011

REPORT DOCUMENTATION PAGE			READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle)  On the Equivalence of Probability Measures		5. TYPE OF REPORT & PERIOD COVERED  Technical	
7. AUTHOR(s)  C.R. Baker & A.F. Gualtierotti		6. PERFORMING ORG. REPORT NUMBER  N00014-75-C-0491 N00014-81-K-0373	
9. PERFORMING ORGANIZATION NAME AND ADDRESS  Department of Statistics University of North Carolina Chapel Hill, North Carolina 27514		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS  NR 042-269 SRO 105	
11. CONTROLLING OFFICE NAME AND ADDRESS  Statistics & Probability Program Office of Naval Research Arlington, VA 22217		12. REPORT DATE	
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		13. NUMBER OF PAGES  10	
		15. SECURITY CLASS. (of this report)  UNCLASSIFIED	
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE	
16. DISTRIBUTION STATEMENT (of this Report)  Approved for public release; distribution unlimited			
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)			
18. SUPPLEMENTARY NOTES			
19. KEY WORDS (Continue on reverse side if necessary and identify by block number)  Signal detection; absolute continuity; likelihood ratio			
20. ABSTRACT (Continue on reverse side if necessary and identify by block number)  Signal detection problems are discussed for the case when the noise IS Gaussian, but signal-plus-noise can be non-Gaussian. A likelihood ratio is given.			

### On the equivalence of probability measures

Let  $\Omega$  be a set ( $\omega$  denotes one of its elements) and  $\mathcal{A}$  be a  $\sigma$ -algebra of subsets of  $\Omega$ . If  $(P, Q)$  is a pair of probability measures on  $\mathcal{A}$ , we are interested in describing the relations which may exist between  $P$  and  $Q$ .

To begin, suppose that  $P$  and  $Q$  are only  $\sigma$ -finite and that  $Q$  is a signed measure.

$Q$  is absolutely continuous with respect to  $P$  (notation:  $Q \ll P$ ) if  $Q(A) = 0$  whenever  $P(A) = 0$  and  $Q$  is singular with respect to  $P$  if there exists a set  $A$  in  $\mathcal{A}$  such that  $Q(A) = P(\Omega \setminus A) = 0$  (notation:  $Q \perp P$ ). When  $Q \ll P$ ,  $Q$  has the following representation:

$$Q(A) = \int_A f(\omega)P(d\omega) ,$$

where  $f$  is measurable with respect to  $\mathcal{A}$  and unique within a set of  $P$ -measure zero.  $f$  is called the Radon-Nikodym derivative of  $Q$  with respect to  $P$  and usually one writes  $dQ/dP$  for  $f$ .

The most general statement about the pair  $(P, Q)$  which is available is the Lebesgue decomposition theorem

$Q = Q_1 + Q_2$  uniquely, where

- (a)  $Q_1$  is a  $\sigma$ -finite signed measure and  $Q_1 \ll P$
- (b)  $Q_2$  is a  $\sigma$ -finite signed measure and  $Q_2 \perp P$

Examples illustrating these concepts are easy to find and we shall give two.

Let:  $\Omega = \mathbb{R}$ ,  $\mathcal{A} = \mathcal{B}(\mathbb{R})$  (the Borel sets of  $\mathbb{R}$ ),  $\text{Leb} \equiv$  Lebesgue measure.

1. Let  $P = \text{Leb}$  and  $Q =$  Gaussian measure with mean zero and variance 1, that is

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

$$Q(A) = \int_A \phi(x)P(dx)$$

Then  $Q \ll P$  and  $dQ/dP = \phi$ . Since  $\phi(x) > 0$ ,

8716

Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	<input type="checkbox"/>
By _____	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

$$P(A) = \int_A P(dx) = \int_A \phi^{-1}(x)\phi(x)P(dx) = \int_A \phi^{-1}(x)Q(dx).$$

Then  $P \ll Q$  and  $dP/dQ = \phi^{-1}$ . When both  $P \ll Q$  and  $Q \ll P$ ,  $P$  and  $Q$  are said to be mutually absolutely continuous (notation  $P \equiv Q$ ).

2. Let

$$\phi_{a,b}(x) = \begin{cases} \frac{1}{b-a} & x \in [a,b] \\ 0 & x \in \mathbb{R} \setminus [a,b] \end{cases} \quad a < b.$$

Set

$$U_{a,b}(A) = \int_A \phi_{a,b}(x)P(dx), \quad P = \text{Leb.}$$

$U_{a,b}$  is called the uniform measure on  $[a,b]$ .  $Q$  is as in 1.

One has

$$U_{a,b}(A) = \int_A \phi_{a,b}(x)\phi^{-1}(x)Q(dx), \quad \text{so that}$$

$$U_{a,b} \ll Q \quad \text{and} \quad dU_{a,b}/dQ = \phi_{a,b} \cdot \phi^{-1}$$

Let  $Q_1(A) = Q(A \cap [a,b])$  and  $Q_2(A) = Q(A \cap (\mathbb{R} \setminus [a,b]))$ . Then, if  $I_{[a,b]}(x)$  is one where  $a \leq x \leq b$  and zero elsewhere,

$$Q_1(A) \equiv \int_A I_{[a,b]}(x)Q(dx) = (b-a) \int_A I_{[a,b]}(x)\phi(x)U_{a,b}(dx)$$

so that  $Q_1 \ll U_{a,b}$  and  $dQ_1/dU_{a,b} = (b-a)I_{[a,b]}(x)\phi(x)$ . Furthermore,  $Q_2([a,b]) = Q(\varphi) = 0$  and  $U_{a,b}(\mathbb{R} \setminus [a,b]) = 0$ , so that  $Q_2 \perp U_{a,b}$ . Since  $Q = Q_1 + Q_2$ ,  $(Q_1, Q_2)$  is the Lebesgue decomposition of  $Q$  with respect to  $U_{a,b}$ .

The problem of obtaining the Lebesgue decomposition is particularly interesting and useful when  $\Omega$  is a space of functions (or of classes of functions) and  $P$  and  $Q$  are induced by stochastic processes. Typically  $\Omega$  is  $C[0,T]$ , the space of continuous functions on  $[0,T]$ , or  $L_2[0,T]$ , the space of equivalence classes of almost surely equal functions on  $[0,T]$  which are square-integrable.  $A$  is then the  $\sigma$ -algebra generated by the open sets of  $\Omega$ . Thus if  $X: \Omega \times [0,T] \rightarrow \mathbb{R}$  is a stochastic process with continuous paths,  $\underline{X}: \Omega \rightarrow C[0,T]$  defined by  $\underline{X}(\omega) = \{X(\omega, t), 0 \leq t \leq T\}$  is a measurable map, so that  $P_X = P \circ \underline{X}^{-1}$  is a probability measure on the Borel sets of  $C[0,T]$ . More generally if  $F$  is a linear space of functions  $f$ ,  $P_X$  is defined by relations of the type:

$$P_X\{f: (\varphi_1(f), \dots, \varphi_n(f)) \in B\} = P\{\omega: (\varphi_1(\underline{X}(\omega)), \dots, \varphi_n(\underline{X}(\omega))) \in B\} ,$$

$\varphi_1, \dots, \varphi_n$  being continuous linear functionals on  $F$ .

There are two major methods to obtain results about the absolute continuity of measures  $P_X$  and  $P_Y$ . The first consists in choosing  $X$  and  $Y$  with probability laws of the same type (two Gaussian, two diffusions, etc.) and using this type to characterize absolute continuity, or equivalence, and the second consists in choosing for  $X$  a martingale (for example Brownian motion) and for  $Y$  a process of the form  $V+X$ , where  $V$  is of bounded variation. The privileged tool is then stochastic calculus. We shall consider below a problem in the second category

and motivate it from considerations which arise in statistical communication theory. If  $V$  is independent of  $X$ , absolute continuity can in certain cases be obtained if absolute continuity is known to hold for  $V$  non-random, without recourse to stochastic calculus.

In statistical communication theory one tries to accurately transmit information over communication channels. The information to be transmitted is called a signal ( $S$ ) and impediments to accurate transmission are called noise ( $N$ ). The theory is statistical in the sense that  $S$  and  $N$  are stochastic processes (impediments can only be known "on average" and information, like the words of a language, has a random component: only the frequency of appearance of given words can be approximated). Communication channels operate as "black boxes": their description from basic physical principles is usually either impossible or too complicated to be of much use. The signal  $S$  is the input to the box and the output is some function of  $S$  and  $N$ , say  $f(S, N)$ . We shall consider here the case of  $f(S, N) = S + N$ , which has proved useful. Most communication systems are monitored continuously and the operator is given a function  $\{x(t), 0 \leq t \leq T\}$  which can be noise or a distorted signal. So, the first aim is to establish whether  $x$  is a realization of  $N$ , that is  $x(t) = N(\omega, t)$ , for some fixed  $\omega$ , or whether  $x$  is a realization of  $S + N$ , that is  $x(t) = N(\omega, t) + S(\omega, t)$ . If a decision can be reached with no possibility of error, the problem is said to be singular:  $F$ , the space of paths for  $N$  and  $S + N$ , can be partitioned into two measurable disjoint subsets  $F_1$  and  $F_2$  such that  $P_N(F_1) = P_{S+N}(F_2) = 1$ . If the requirement that a correct decision be always reached, say in case of no signal, forces the user to ignore the data, the problem is said to be non-singular: for every measurable subset  $F_1$  of  $F$ ,  $P_N(F_1) = 1$  implies  $P_{S+N}(F_1) = 1$ . Singularity corresponds thus to the case  $P_N \perp P_{S+N}$  and non-singularity to the

case  $P_N = P_{S+N}$ . In the latter case, a good decision procedure consists in looking at  $dP_{S+N}/dP_N(x)$  and in deciding that a signal is present when this is large (that such a procedure is good follows from the Neyman-Pearson lemma of statistics).  $dP_{S+N}/dP_N$  can always be thought of as the ratio of two densities, say  $D_{S+N}$  and  $D_N$

$$P_{S+N}(A) = \int_A \left[ \frac{dP_{S+N}}{d[P_N + P_{S+N}]} \Bigg/ \frac{dP_N}{d[P_N + P_{S+N}]} \right] d[P_N + P_{S+N}] .$$

and the choice consists in deciding "no signal" if  $cD_N(x) \geq D_{S+N}(x)$  and "signal" if  $cD_N(x) < D_{S+N}(x)$ . (for some adequately chosen c).

A classical problem of the type just described has the form:  $N$  is a Gaussian process, but no frequency requirements are imposed on  $S$ . The results which shall be stated have been obtained by C. R. Baker and A. F. Gualtierotti (UT Electrical Engineering, Fall 1981).

#### The model:

$(\Omega, \mathcal{A}, P)$  is the basic probability space specified by the experiment and the information as time evolves is contained in  $\underline{\mathcal{A}} = \{\mathcal{A}_t, 0 \leq t \leq T\}$ , where  $\mathcal{A}_t$  is a  $\sigma$ -algebra of subsets of  $\Omega$ , and  $\mathcal{A}_t \subseteq \mathcal{A}$ ,  $s < t \Rightarrow \mathcal{A}_s \subseteq \mathcal{A}_t$ .

#### (A) The noise $N$

Let  $\underline{B}(\omega, t) = (B_1(\omega, t), \dots, B_n(\omega, t))^t$ , the processes  $B_k$  being independent, Gaussian with independent increments and with continuous variance  $B_k(t) = E B_k^2(\omega, t)$  ( $E B_k(\omega, t) = 0$ ).  $\underline{B}$  is adapted to  $\underline{\mathcal{A}}$ , that is  $\underline{B}(\cdot, t)$  is measurable with respect to  $\mathcal{A}_t$ . The diagonal matrix with entries  $B_k(t)$  is denoted  $\Sigma_B(t)$ . It can be shown that  $B_k$  is a continuous square integrable martingale with increasing process  $B_k$  (one can choose for  $B_k$  the Wiener process  $W_k$  and then  $B_k(t) = t$ ).  $F(t, x)$  is a function defined on  $[0, T] \times [0, T]$ : it is measurable and  $F(t, x) = 0$  for  $x > t$ .

Furthermore, it is assumed that  $\int_0^T \langle \underline{F}(t, x), \Sigma_B(dx) \underline{F}(t, x) \rangle$  is finite for all  $t$ .

$N$  is then defined by

$$N(\omega, t) = \int_0^T \langle \underline{F}(t, x), \underline{B}(\omega, dx) \rangle$$

and is supposed mean square continuous, that is

$$E[N(\cdot, t) - N(\cdot, s)]^2 = \int_0^T \langle \underline{F}(t, x) - \underline{F}(s, x), \Sigma_B(dx) [\underline{F}(t, x) - \underline{F}(s, x)] \rangle$$

is small when  $|t-s|$  is.

Remark: This choice of  $N$  is based on the Cramer-Hida representation: a second order, mean square continuous stochastic process  $X$  can always be written in the form

$$X(\omega, t) = \int_0^t \langle \underline{F}(t, x), \underline{B}(\omega, dx) \rangle$$

where the  $B_k$ 's have orthogonal increments. If  $X$  is Gaussian,  $\underline{B}$  is Gaussian. However  $\underline{B}$  can have an infinite number of components.

(B) The signal  $S$

$$S(\omega, t) = \int_0^t \langle \underline{F}(t, x), \Sigma_B(dx) \underline{s}(\omega, x) \rangle$$

$\underline{s}$  is assumed optional, which is a regularity condition imposed by the recourse to stochastic integration (the prototype of an optional process is a process whose trajectories are functions  $f$  which are continuous to the right and have limits to the left). A further assumption is that

$$E \int_0^T \langle s(\omega, x), \sum_B (dx) s(\omega, x) \rangle < \infty$$

Remark: Let  $R_N(s, t) = EN(\cdot, s)N(\cdot, t) = \int_0^T \langle F(t, x), \sum_B (dx) F(t, x) \rangle$ . With  $R_N$ , one can associate a Hilbert space of functions  $H(R_N)$  in the following way: on the vector space of functions  $\sum_{k=1}^n a_k R_N(\cdot, t_k)$ , define

$$\left\langle \sum_{k=1}^n a_k R_N(\cdot, t_k), \sum_{l=1}^p b_l R_N(\cdot, s_l) \right\rangle_{H(R_N)} = \sum_{k=1}^n \sum_{l=1}^p a_k b_l R_N(s_l, t_k) .$$

$H(R_N)$  is the completion of this vector space for the norm  $\|\cdot\|_{H(R_N)}$  and is called the reproducing kernel Hilbert space of  $R_N$ . The name comes from the relation

$$\langle R_N(\cdot, x), \sum_{k=1}^n a_k R_N(\cdot, t_k) \rangle_{H(R_N)} = \sum_{k=1}^n a_k R_N(x, t_k) .$$

Since

$$\sum_{k=1}^n \sum_{l=1}^p a_k b_l R_N(s_l, t_k) = \int_0^T \left\langle \sum_{k=1}^n a_k F(t_k, x), \sum_B (dx) \sum_{l=1}^p b_l F(s_l, x) \right\rangle$$

and

$$\sum_{k=1}^n a_k R_N(u, t_k) = \int_0^T \left\langle F(u, x), \sum_B (dx) \sum_{k=1}^n a_k F(t_k, x) \right\rangle$$

one has here a representation of  $H(R_N)$ :

$$H(R_N) = \{a(t) = \int_0^T \langle F(t, x), \sum_B (dx) A(x) \rangle\}$$

Ae subspace of  $L_2[\Sigma_B]$  generated by  $F(t, \cdot)$

$$\langle a, \tilde{a} \rangle_{H(R_N)} = \int_0^T \langle A(x), \sum g(dx) \tilde{A}(x) \rangle$$

One thus sees that the signal  $S$  has been chosen to belong to the reproducing kernel Hilbert space of the noise. This is no coincidence: in many instances this is a necessary condition for equivalence.

Example: Let  $N = W$ , a Wiener process on  $[0, T]$ . Then

$$H(R_N) = \{a(t) = \int_0^t a(x) dx, \int_0^T a^2(x) dx < \infty\}$$

$$\langle a, \tilde{a} \rangle_{H(R_N)} = \int_0^T a(x) \tilde{a}(x) dx = \langle a, \tilde{a} \rangle_{L_2[0, T]}$$

It can be shown that the process  $W_a(\omega, t) = a(t) + W(\omega, t)$  has a law  $P_{W_a}$  equivalent to  $P_W$ , the law of  $W$ , if and only if  $a \in H(R_N)$ . More generally, the law  $P_Y$  of the process  $Y$  is equivalent to  $P_W$  "if and only if"  $Y$  has the form

$$Y(\omega, t) = \int_0^t a(\omega, x) dx + W(\omega, x)$$

with

$$P \left\{ \int_0^T a^2(\omega, x) dx < \infty \right\} = 1$$

(The " " indicate that many details concerning  $a$  and  $Y$  have been left out.)

We now have the following results:

- I) There exists a measurable map  $T: C_n[0, T] \rightarrow L_2[0, T]$  ( $C_n[0, T] = \{f : [0, T] \rightarrow \mathbb{R}^n, f \text{ is continuous}\}$ ) such that  $N = T \circ \underline{B}$  and  $S + N = T \circ \underline{Z}$ , where  $\underline{Z}(\omega, t) = \int_0^t \sum g(dx) \underline{s}(\omega, x) + \underline{B}(\omega, t)$ , provided  $P_Z \ll P_B$ .

II) If I) obtains,  $P_{S+N} \ll P_N$  and

$$\frac{dP_{S+N}/dP_N}{dP_Z/dP_B}(\underline{h}) = \int_{C_n[0,T]} dP_Z/dP_B(\underline{c}) P_{B/N=\underline{h}}(d\underline{c})$$

where  $P_{B/N=\underline{h}}$  is the conditional law of  $\underline{B}$  knowing that  $N = \underline{h}$ .

III)  $P_{B/N=\underline{h}}$  is a point mass concentrated at  $\underline{m}(t, \underline{h})$ , with

$$m_1(t, \underline{h}) = \sum_{k=1}^{\infty} \frac{\langle f_i(t, \cdot), \psi_k \rangle_{L_2[0,T]} \langle \underline{h}, \psi_k \rangle_{L_2[0,T]}}{\lambda_k}$$

where  $(\lambda_k, \psi_k)$  is an eigenvalue-eigenvector couple for the operator  $R^N$  on  $L_2[0,T]$  defined by  $(R^N f)(t) = \int_0^T R_N(t,x) f(x) dx$ .  $N$  being mean square continuous,  $R_N$  is continuous and thus  $R^N$  has finite trace:

$$R^N = \sum_{k=1}^{\infty} \lambda_k \psi_k \otimes \psi_k \quad ([a \otimes b](x) = \langle x, b \rangle_{L_2[0,T]} a)$$

The function  $f_i(t, \cdot)$  is given by  $f_i(t, x) = \int_0^t F_i(x, u) \beta_i(du)$ . Hence  $P_N$  - almost surely,

$$\frac{dP_{S+N}/dP_N}{dP_Z/dP_B}(\underline{h}) \approx dP_Z/dP_B(\underline{m}(\cdot, \underline{h}))$$

One need thus define explicitly the function on  $C_n[0,T]$ ,  $dP_Z/dP_B$ .

Let  $\Pi(\underline{c}, t) = \underline{c}(t)$ ,  $\underline{c} \in C_n[0,T]$ . It can be shown that if  $\underline{Z}$  is the solution of the stochastic differential equation

$$\underline{Z}(\omega, t) = \int_0^t \Sigma_B(dx) \underline{a}(\underline{Z}(\omega, \cdot), x) + \underline{B}(\omega, t) ,$$

where  $\underline{\alpha}(Z(\omega, \cdot), x)$  is a function of  $Z(\omega, u)$ , for  $u \leq x$ , then,  $P_B$  - almost surely

$$dP_Z/dP_B(c) = \exp \left\{ \int_0^T \langle \underline{\alpha}(c, x), \Pi(c, dx) \rangle - \frac{1}{2} \int_0^T \langle \underline{\alpha}(c, x), \Sigma_B(dx) \underline{\alpha}(c, x) \rangle \right\}$$

The term  $\int_0^T \langle \underline{\alpha}(c, x), \Pi(c, dx) \rangle$  is a stochastic integral with respect to the process  $\{\Pi(\cdot, t), t \in [0, T]\}$  defined on the probability space  $(C_n[0, 1], \mathcal{B}(C_n[0, 1]), P_B)$ .

Since  $P_B$  and  $P_{B/N=h}$  are orthogonal, one cannot simply evaluate the exponential containing the stochastic integral at the point  $\underline{m}(\cdot, h)$ . However, one can approximate  $dP_Z/dP_B(\underline{m}(\cdot, h))$  in  $L_1[P_N]$  sense, replacing  $dP_Z/dP_B$  by any  $L_1[P_B]$ -approximation, which is useful for practical applications.

The explicit representation of  $dP_Z/dP_B$  is due to two facts: one is Girsanov's theorem, which gives conditions under which the translation of a martingale by a process of bounded variation is again a martingale with the same local characteristics (in fact, in its generality, the theorem says that the class of semimartingales, that is sums of martingales and processes of bounded variation, is invariant under absolutely continuous changes of probability measures), and the other is the characterization of Brownian motion as a continuous martingale. Since these two facts are valid for example for processes with independent increments, the results just stated are valid for processes other than the Gaussian ones. Finally, it is also possible to have a  $B$  with an infinite number of components.